

Optimal Locally Repairable Codes with Improved Update Complexity

Mehrtash Mehrabi,

Mostafa Shahabinejad,

Masoud Ardakani, and

Majid Khabbazzian

Abstract—For a systematic erasure code, update complexity (UC) is defined as the maximum number of parity blocks needed to be changed when some information blocks are updated. Locally repairable codes (LRCs) have been recently proposed and used in real-world distributed storage systems. In this paper, update complexity of optimal LRCs is studied and both lower and upper bounds on UC are established in terms of length (n), dimension (k), minimum distance (d), and locality (r) of the code, when $(r+1) \mid n$. Furthermore, a class of optimal LRCs with small UC is proposed. Our proposed LRCs could be of interest as they improve UC without sacrificing optimality of the code.

Index Terms—Update complexity, erasure coding, distributed storage system, locally repairable codes.

I. INTRODUCTION

Distributed storage systems (DSSs) are used to store large-scale data in a secure and reliable way. A DSS uses a number of storage nodes, called data nodes, to store data. In such storage systems, data node failures occur frequently due to several reasons such as hardware/software problems associated with the underlying network or data nodes. In order to recover the lost/erased data, redundancy is required. For example, the approach of keeping several replicas of data in distinct data nodes, known as replication, is widely used [1]. The high storage overhead of the replication method results in a costly maintenance for DSSs [2].

Recently, systematic erasure codes have been proposed and used in DSSs to decrease storage overhead. In order to use an erasure code in a DSS, first, a stripe of data is split into k information blocks. Then, using an (n, k) erasure code, n encoded blocks are generated from k information blocks. The n encoded blocks are then stored in n different data nodes. Hence, for an (n, k) erasure code with minimum distance d , where $d \leq n - k + 1$, $n - k$ parity blocks are generated such that k information blocks can be recovered by any $n - d + 1$ encoded blocks. Such, systematic linear block erasure codes¹ have been used in real-world cloud storage systems such as Google File System [3], Microsoft Azure Storage [2], and Facebook HDFS-RAID [1].

Although erasure codes reduce storage overhead, they need access to many nodes to recover a missing data block.

¹In systematic codes, information blocks can be directly stored and read with no encoding and decoding processes [1], [2]. This is why, in DSSs, systematic codes are preferred to the non-systematic ones. From now on, whenever we say any class of erasure codes we mean the systematic erasure codes of that class.

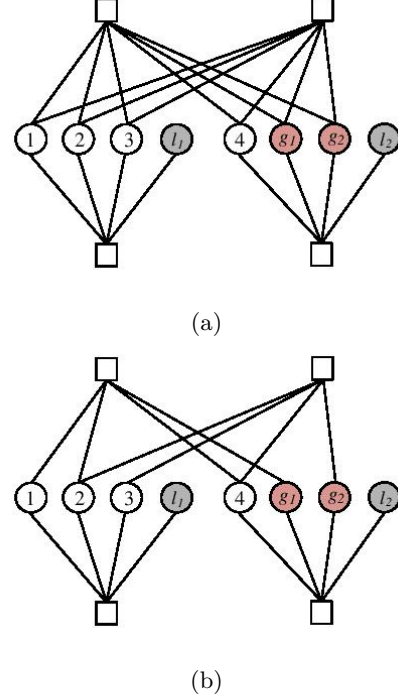


Figure 1: Tanner graphs of two optimal $(n, k, d, r) = (8, 4, 4, 3)$ LRCs with different values of UC. The average UCs associated with Figs. 1a and 1b are $\bar{u}_1 = (3 \times 4 + 1 \times 3)/4 = 3.75$ and $\bar{u}_2 = (3 \times 3 + 1 \times 4)/4 = 3.25$, respectively.

For example, an (n, k) erasure code with $d = n - k + 1$, known as maximum distance separable (MDS) codes, requires k data nodes in order to recover one missing data block.

Decreasing the number of participating nodes in a recovery process is crucial as it decreases the costly repair bandwidth and disk I/O. The maximum number of active nodes required to recover a failed node, denoted r , is defined as the locality of an erasure code. Locally repairable codes (LRCs) are a class of codes that are designed for small r . The following bound is obtained for minimum distance of LRCs [4], [5]

$$d \leq n - k - \left\lceil \frac{k}{r} \right\rceil + 2. \quad (1)$$

LRCs that achieve this bound are called *optimal*. It is verified that the bound in (1) is tight if $(r+1) \mid n$ [5]. From now on, we assume that $(r+1) \mid n$.

Recently, there have been a lot of studies on LRCs.

For example, in [6], a tight upper bound on d for LRCs with $\frac{n}{r} > \lceil \frac{n}{r+1} \rceil$ is proposed, and LRCs with the largest possible d with $r \leq \sqrt{n} - 1$ are designed. In [7], LRCs with minimum storage regenerating and minimum bandwidth regenerating are considered. In [8]–[10], LRCs over small fields are proposed in order to decrease the computational complexity associated with coding.

For a systematic (n, k, d, r) optimal LRC, there exist $n - k$ parity blocks constructed from k information blocks. While some of these $n - k$ parity blocks are constructed locally from a few blocks to achieve the code locality, some other parity blocks are constructed globally to achieve the required minimum distance. In the existing optimal LRCs, all the information blocks get involved in these globally constructed parity blocks [11]–[13]. Consequently, if some information blocks have to be updated, all the global parity blocks have to be changed resulting in a costly update process.

Is it possible to generate different optimal (n, k, d, r) LRCs with different update complexity? If yes, how can we find optimal LRCs with small update complexity? This is the central question studied in this paper.

Fig. 1 shows the Tanner graphs of two optimal $(n, k, d, r) = (8, 4, 4, 3)$ LRCs. In the LRC of Fig. 1a, all the information blocks are involved in the two parity blocks g_1 and g_2 . However, in the LRC of Fig. 1b, only some information blocks are involved in g_1 and g_2 . In this example, if one information block is updated, the LRC of Fig. 1b, on average, needs 13% less updates on the parity blocks.

In this paper, we consider the problem of update complexity (UC) for systematic optimal LRCs. The contributions of this paper are two-folds. By taking an existing definition of update complexity and generalizing it, we obtain both the upper and lower bounds on UC for an importance class of LRCs. Furthermore, we design a class of *optimal* LRCs whose average UC is close to the obtained lower bound. Note that this improvement of UC is achieved without sacrificing other important parameters of the LRC such as minimum distance (d), rate ($\frac{k}{n}$), or locality (r).

The remainder of this paper is organized as follows. In Section II, we provide the required preliminaries. In Section III, we obtain lower and upper bounds on UC. In Section IV, we introduce our proposed LRC with small UC and we compare our LRC with other LRCs in terms of UC. Finally, in Section V, we conclude the paper.

Notations: We show matrices and vectors by capital bold letters and bold letters, respectively. \mathbb{F}_q and \otimes stand for a finite field of order q and tensor product, respectively. \mathbf{I}_a and $\mathbf{0}_{b \times c}$ represent an identity matrix of size a and a zero matrix of size $b \times c$, respectively. $(\cdot)^T$ and $\mathbf{1}_a$ represent matrix transpose operation and a column vector of ones with size a , respectively. \bar{u} represents $\frac{1}{n} \sum_{i=1}^n u_i$ which is the average of u_i 's for $i \in \{1, \dots, n\}$. For an integer n , $[n] = \{1, \dots, n\}$.

II. PRELIMINARIES

A. Definitions

Systematic linear block codes: The generator matrix of an (n, k) systematic linear block code can be presented as $\mathbf{G} = [\mathbf{I}_k, \mathbf{P}] \in \mathbb{F}_q^{k \times n}$, where $\mathbf{P} \in \mathbb{F}_q^{k \times (n-k)}$. Assuming that $\mathbf{x} = [x_1, x_2, \dots, x_k] \in \mathbb{F}_q^{1 \times k}$ and $\mathbf{y} = [y_1, y_2, \dots, y_n] \in \mathbb{F}_q^{1 \times n}$ are the information and encoded vectors, respectively, we have $\mathbf{y} = \mathbf{x}\mathbf{G}$. The parity check matrix of the code is $\mathbf{H} = [-\mathbf{P}^T, \mathbf{I}_{n-k}] \in \mathbb{F}_q^{(n-k) \times n}$ satisfying $\mathbf{G}\mathbf{H}^T = \mathbf{0}_{1 \times (n-k)}$.

Remark 1. In a DSS, in order to store a stripe of data of size L symbols by an (n, k) systematic linear block code, first, the stripe is partitioned into k data blocks each of size $l = \frac{L}{k}$ symbols. Assume that $x_{i,j}$ is i -th symbol of j -th data block, where $i \in [l]$ and $j \in [k]$. Then, $\mathbf{x}_i = [x_{i,1}, \dots, x_{i,k}] \in \mathbb{F}_q^{1 \times k}$. The coded vector $\mathbf{y}_i = [y_{i,1}, \dots, y_{i,n}] \in \mathbb{F}_q^{1 \times n}$ is generated as $\mathbf{y}_i = \mathbf{x}_i\mathbf{G} = \mathbf{x}_i[\mathbf{I}_k, \mathbf{P}]$. From which, matrix $\mathbf{Y} \in \mathbb{F}_q^{l \times n}$ is constructed by stacking l encoded vectors \mathbf{y}_i . Each column of \mathbf{Y} is an encoded block which is stored in a data node. For simplicity, from now on, we assume that $l = 1$.

Minimum distance of code (d): The minimum Hamming distance between any two codewords of an erasure code is defined as the minimum distance of that code, denoted d . Any (n, k) erasure code with minimum distance d tolerates any $d - 1$ symbol erasures.

Locality: For an (n, k) linear block code, locality of the i -th encoded block where $i \in [n]$, denoted r_i , is defined as the minimum number of other blocks needed to participate in its recovery process. In other words, in the case that the i -th block is missing, at least r_i other blocks are needed to reconstruct it. Locality of a code, denoted r , is defined as the maximum of r_i for $i \in [n]$, i.e. $r = \max_{i \in [n]} r_i$.

Tanner/Factor graph: A Tanner graph (Factor graph) [14], [15] associated with an (n, k) linear block code is a bipartite graph with n variable nodes on one side (usually shown by circles) and $n - k$ check nodes on the other side (usually shown by squares) which represent columns and rows of the code's parity check matrix $\mathbf{H} \in \mathbb{F}_q^{(n-k) \times n}$, respectively. There is an edge between j -th ($j \in [n]$) variable node and i -th ($i \in [n - k]$) check node in the Tanner graph, if $h_{i,j}$ is non-zero, where $h_{i,j}$ represent the element of i -th row and j -th column of \mathbf{H} .

Information and parity nodes: Consider the Tanner graph associated with a systematic (n, k) linear block code. Among all the n variable nodes, k variable nodes are corresponding with k information blocks. We call these variable nodes *information nodes* and represent them by white circles in the Tanner graph. The remaining $n - k$ variable nodes are corresponding with $n - k$ parity blocks. We call these variable nodes *parity nodes* and represent them by shaded circles in the Tanner graph (see Fig. 1 as an example).

Local and mixed groups, local and global check nodes and parity nodes: In the Tanner graph of an (n, k, d, r)

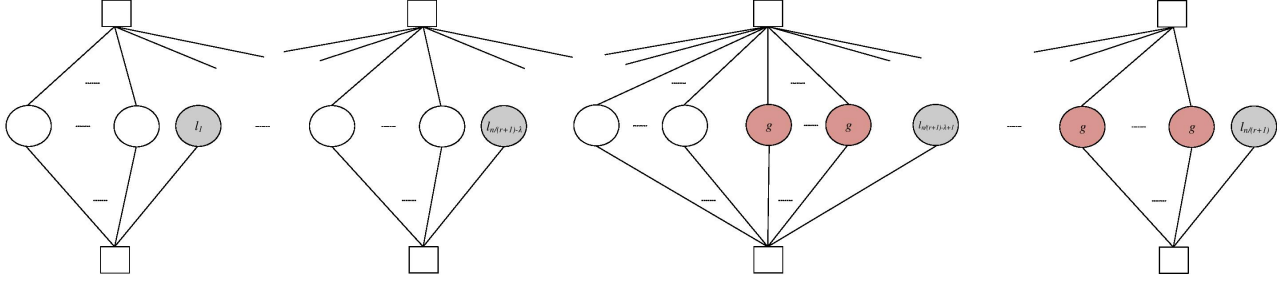


Figure 2: Construction of an (n, k, d, r) ν -LRC. There are $\frac{n}{r+1}$ local and $d - 2 - \lfloor \frac{d-2}{r+1} \rfloor$ global parity nodes, where $r + 1$ is cardinality of each local group.

LRC, among $n - k$ check nodes, a minimal set of check nodes, each having at most $r + 1$ edges that cover all variable nodes are called *local check nodes* and constitute *local groups*. In other words, a failed variable node can be reconstructed within its local group. Note that the locality r_i of each variable is the size of its local group minus one. Check nodes which are not local are called *global check nodes*. Furthermore, parity nodes associated with local check nodes are called *local parity nodes*; and the rest of parity nodes associated with global check nodes are called *global parity nodes*. Local groups containing global parity nodes are called *mixed groups*. For example, in Figs. 1a and 1b, the check nodes that are below the variable nodes are local check nodes and the rest are global. In this example, the local group containing the fourth information node is a mixed group.

B. An Important Class of Optimal LRCs

Non-Overlapped and Uniform optimal LRCs (ν -LRCs): An important class of LRCs are optimal LRCs where $(r + 1) \mid n$. The importance of this class stems from the fact that when $(r + 1) \mid n$ the bound in (1) can be achieved with equality. In fact, such optimal LRCs have been the focus of studies in [11]–[13], [16]. Hence, we also assume $(r + 1) \mid n$.

Assuming that $(r + 1) \mid n$, ν -LRCs are a class of optimal LRCs in which n encoded blocks are partitioned uniformly into $\frac{n}{r+1}$ non-overlapped local groups, where $r + 1$ is the cardinality of each local group. In the structure of ν -LRCs, local groups are both uniform and non-overlapped, hence the name. In ν -LRCs, among the total $(n - k)$ check nodes, there are $\frac{n}{r+1}$ local and $n - k - \frac{n}{r+1}$ global check nodes (see Fig. 2 as an example).

In the following remark, the exact number of global check nodes for an (n, k, d, r) ν -LRC is computed.

Remark 2. Since ν -LRC are optimal, by (1), the number of their global check nodes can also be expressed as

$$n - k - \frac{n}{r + 1} = d - 2 - \left\lfloor \frac{d - 2}{r + 1} \right\rfloor.$$

III. BOUND ON UPDATE COMPLEXITY (UC)

In this section, we formally define UC and then for the class of LRC that we discussed in the previous section, we find upper and lower bounds on UC.

In [17], update complexity (UC) is defined as the maximum number of parity blocks needed to be changed when an information block is updated. By generalizing this definition, we define UC, as the number of parity blocks needed to be changed when a set of x , $x \in [k]$, information blocks are updated, and denoted by u_x .

In this paper, we study the UC of ν -LRCs and seek ν -LRCs that have low UC. In the following, we start with studying a special variable node arrangement for ν -LRCs. Later, we will show how this arrangement helps us to establish our bounds.

Fig. 2 shows the general structure of an (n, k, d, r) ν -LRCs with a special variable node arrangement. Here, n encoded nodes are partitioned into $\frac{n}{r+1}$ local groups of size $r + 1$. Each local group has a local parity node, denoted by l_i , where $i \in [\frac{n}{r+1}]$. Also, each local group except mixed groups has r information nodes. In general, the number of mixed groups can be arbitrarily large. However, in the structure shown in Fig. 2, the number of mixed group is minimal. This can lead to a small UC. Also, there is at most one mixed group containing both information and global parity nodes. We call this group *infomixed group*². As stated in Remark 2, there are $d - 2 - \lfloor \frac{d-2}{r+1} \rfloor$ global check nodes. Hence, the total number of mixed groups, denoted λ , is

$$\lambda = \left\lceil \frac{d - 2 - \lfloor \frac{d-2}{r+1} \rfloor}{r} \right\rceil.$$

Consequently, the infomixed group, if exists, has $d - 2 - \lfloor \frac{d-2}{r+1} \rfloor - (\lambda - 1)r$ global parity nodes and one local parity node. Thus, there are $r + 1 - (d - 2 - \lfloor \frac{d-2}{r+1} \rfloor - (\lambda - 1)r) - 1$ information nodes in the infomixed group. Observe that the total number of information nodes is

$$k = (\frac{n}{r + 1} - \lambda)r + (r - d + 2 + \left\lfloor \frac{d - 2}{r + 1} \right\rfloor) + (\lambda - 1)r.$$

where the first term is the number of information nodes in none-mixed groups and the second term is that in

²There is no infomixed group, if $r \mid k$.

infomixed group. Now by using the given properties of ν -LRCs, in the following theorem, we establish both lower and upper bounds on UC.

Theorem 1. For an (n, k, d, r) ν -LRC, UC associated with changing x information blocks (u_x) is bounded as

$$d - 1 + \left\lceil \frac{x - (r\lambda - (d - 2 - \lfloor \frac{d-2}{r+1} \rfloor))}{r} \right\rceil \leq u_x \leq d - 1 + \theta,$$

where $\lambda = \left\lceil \frac{d-2 - \lfloor \frac{d-2}{r+1} \rfloor}{r} \right\rceil$, and

$$\theta = \begin{cases} x & \text{if } x \leq \frac{n}{r+1} - \lambda \\ \frac{n}{r+1} - \lambda & \text{otherwise} \end{cases}.$$

Proof: Please refer to Appendix A. ■

Update complexity associated with a set of x information blocks (u_x), can vary from one set to another set, both with cardinality x . Hence, another measure of the update complexity of a code is needed. We use the average UC as a metric to evaluate UC of a code. Average UC of code, denoted $\overline{u_x}$, is defined as the average number of parity blocks needed to be changed when any set of x information blocks, where $x \in [k]$, are updated. In this work, our focus is on $\overline{u_1}$, which is the average update complexity when only one information block is changed. Observe that $\overline{u_x}$ can be enhanced by improving $\overline{u_1}$. In the following theorem, both lower and upper bounds on $\overline{u_1}$ is computed.

Theorem 2. For an (n, k, d, r) ν -LRC, average UC associated with changing one information block ($\overline{u_1}$) is bounded as

$$(d - 1) + \frac{\eta \left\lfloor \frac{\frac{n}{r+1} - \lambda}{\lfloor \frac{d-2}{r+1} \rfloor + 1} \right\rfloor}{k} \leq \overline{u_1} \leq d$$

where $\lambda = \left\lceil \frac{d-2 - \lfloor \frac{d-2}{r+1} \rfloor}{r} \right\rceil$,

$$\eta = \begin{cases} 0 & \text{if } \frac{\alpha}{(d-2 - \lfloor \frac{d-2}{r+1} \rfloor)} \mid \beta \text{ or } \lfloor \frac{d-2}{r+1} \rfloor = \lceil \frac{d-2}{r+1} \rceil \\ \alpha - \beta \lfloor \frac{\alpha}{\beta} \rfloor & \text{otherwise} \end{cases},$$

$\alpha = ((r+1)(\lfloor \frac{d-2}{r+1} \rfloor + 1) - (d-2))(d-2 - \lfloor \frac{d-2}{r+1} \rfloor)$, and $\beta = r(\lfloor \frac{d-2}{r+1} \rfloor + 1)$.

Proof: Please refer to Appendix B. ■

In the following section, we propose a class of ν -LRCs with small average UC, close or even in some cases equal to the lower bound on $\overline{u_1}$.

IV. OUR PROPOSED OPTIMAL LRCs WITH SMALL UC

Here, we present our proposed LRCs using their Tanner graphs. Our proposed LRCs achieve the bound in (1), i.e. they are optimal. Furthermore, they benefit from a small $\overline{u_1}$, close or even equal to the lower bound obtained in Theorem 2. In other words, in comparison with the existing optimal LRCs, our proposed LRCs require accessing and changing a smaller number of parity blocks in the case of information block updates.

A. Construction of Our Proposed Optimal LRCs

In order to construct our proposed LRCs, first, n variable nodes are partitioned into $\frac{n}{r+1}$ local groups each containing $r+1$ variable nodes. Hence, there are $\frac{n}{r+1}$ local check nodes associated with $\frac{n}{r+1}$ local groups, where each local group constructs one local parity block. The remaining $n - k - (\frac{n}{r+1}) = d - 2 - \lfloor \frac{d-2}{r+1} \rfloor$ check nodes construct $d - 2 - \lfloor \frac{d-2}{r+1} \rfloor$ global parity blocks which are placed in the mixed groups (Fig. 2).

By Theorem 1, we have

$$d - 1 \leq u_1 \leq d.$$

This implies that updating an information block requires updating either $d - 1$ or d parity blocks. Therefore, to minimize $\overline{u_1}$, we have to find tanner graphs with minimum number of information nodes whose update requires changing d parity nodes. While constructing such tanner graphs, we have to ensure the minimum distance constraint is satisfied, and for that we use the following theorem from [19].

Theorem 3. [19] There is an erasure code with minimum distance d associated to Tanner graph \mathcal{T} iff every γ check nodes of \mathcal{T} cover $\gamma + k$ variable nodes, where $\gamma \in [n - k - d + 2, n - k]$.

By Theorem 3, a necessary condition to guarantee the minimum distance d for our proposed LRCs is that any collection of $n - k - d + 2$ check nodes consisting of $n - k - d + 1$ local check nodes and a single global check node must cover at least $n - (d - 2)$ variable nodes. The number of local groups outside the selected collection is $\frac{n}{r+1} - (n - k - d + 1) = \lfloor \frac{d-2}{r+1} \rfloor + 1$. To satisfy this condition, the single global check node in any such collections must be connected to all the variable nodes of the local groups outside the collection with at most $d - 2$ exceptions. In other words, at least $(r+1)(\lfloor \frac{d-2}{r+1} \rfloor + 1) - (d - 2)$ variable nodes of any set of $\lfloor \frac{d-2}{r+1} \rfloor + 1$ local groups have to be connected to each of the $d - 2 - \lfloor \frac{d-2}{r+1} \rfloor$ global check nodes.

Considering this, our proposed LRCs can be constructed using Algorithm 1 presented in the next page. Note that by choosing coefficients of parity check matrix associated with the obtained Tanner graph randomly from a sufficiently large Galois field, the LRC can be generated.

B. Properties and Evaluation of Our Proposed Optimal LRCs

In the following, we verify some important properties of our proposed LRCs generated by Algorithm 1.

Proposition 4. The proposed optimal (n, k, d, r) LRCs constructed by Algorithm 1 have minimum distance d .

Proof: Please refer to Appendix C. ■

Remark 3. Our proposed LRC is optimal since all the assumptions in the construction of our proposed LRCs are made based on the satisfaction of bound in (1).

Algorithm 1 Construction of optimal LRCs with small UC

First, construct local groups based on the structure of a ν -LRC depicted in Fig. 2 **then**

- Connect each of the $d - 2 - \lfloor \frac{d-2}{r+1} \rfloor$ global check node to a distinct global parity node located in the mixed groups.
- Connect all the information nodes of the infomixed group, if exists, to all the $d - 2 - \lfloor \frac{d-2}{r+1} \rfloor$ global check nodes.

then

for $i \in [d - 2 - \lfloor \frac{d-2}{r+1} \rfloor]$, **do**

while There is a set of $\lfloor \frac{d-2}{r+1} \rfloor + 1$ local groups out of all $\frac{n}{r+1}$ local groups, where i -th global check node is not connected to at least $(r+1)(\lfloor \frac{d-2}{r+1} \rfloor + 1) - (d-2)$ variable nodes of that set **do**

Connect the i -th global check node to $(r + 1)(\lfloor \frac{d-2}{r+1} \rfloor + 1) - (d - 2)$ information nodes of the selected set.

end

end

if There is an information node in the non-mixed groups connected to β global check nodes, where $\beta < (d - 2 - \lfloor \frac{d-2}{r+1} \rfloor - (\lambda - \lfloor \frac{d-2}{r+1} \rfloor))$ and $\lambda = \lfloor \frac{d-2 - \lfloor \frac{d-2}{r+1} \rfloor}{r} \rfloor$ **then**

connect it to $d - 2 - \lambda - \beta$ global check nodes in order to have connection with at least $d - 2 - \lfloor \frac{d-2}{r+1} \rfloor - (\lambda - \lfloor \frac{d-2}{r+1} \rfloor)$ global check nodes

end

Remark 4. Our proposed optimal LRCs improve the UC compared to the existing solutions. The construction of our proposed LRCs ensures that not all information nodes be involved in global check nodes. In fact, it tries to keep the number of information nodes involved in any given global check node small. This means a small number of global parity blocks need update when information blocks are updated. This in turn results in codes with average UC close to the lower bound.

Remark 5. In the case that all information nodes in the non-mixed groups have the same degree and each of them is connected to exactly $d - 2 - \lfloor \frac{d-2}{r+1} \rfloor - (\lambda - \lfloor \frac{d-2}{r+1} \rfloor)$ global check nodes, the lower bound of single updates ($u_1 = d - 1$) for all information nodes is achieved. For our proposed optimal (n, k, d, r) LRC, this is the case when $((r+1)(\lfloor \frac{d-2}{r+1} \rfloor + 1) - (d - 2)) \mid \beta$, where $\beta = r(\lfloor \frac{d-2}{r+1} \rfloor + 1)$, is satisfied.

Example 1. Tanner graph of an optimal $(n, k, d, r) = (15, 9, 5, 4)$ LRC is obtained using the Algorithm 1 (see Fig. 3). In this example, we have

$$((r+1)(\lfloor \frac{d-2}{r+1} \rfloor + 1) - (d - 2)) = 2 \mid \beta = 4.$$

Hence, Remark 5 is satisfied and all the information nodes are connected to $d - 1 = 4$ parity nodes.

In order to evaluate UC of our proposed coding scheme, we compare our proposed $(n, k, d, r) = (15, 9, 5, 4)$ optimal

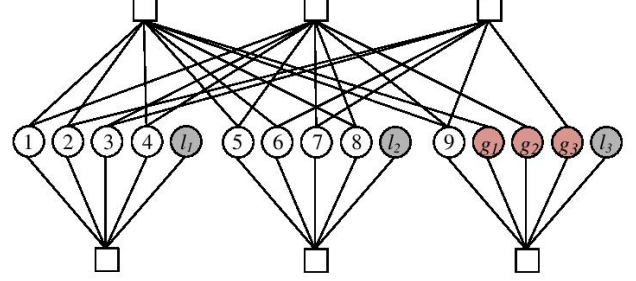


Figure 3: Tanner graph of an $(n, k, d, r) = (15, 9, 5, 4)$ optimal LRC. In this figure, the global check nodes are connected to variable nodes based on our proposed method.

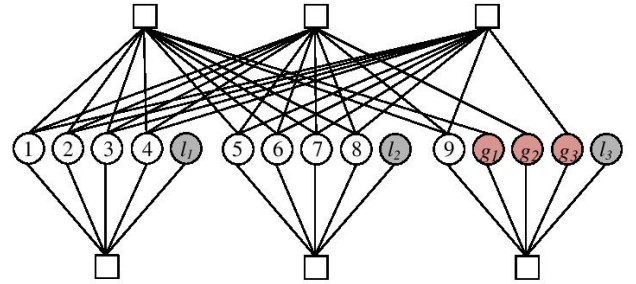


Figure 4: Tanner graph of an $(n, k, d, r) = (15, 9, 5, 4)$ optimal LRC. In this figure, each of the global check node are connected to $k + 1 = 10$ variable nodes.

LRC presented in Example 1 with one whose global check nodes are connected to $k + 1$ variable nodes, which is the common approach in designing optimal LRCs (please see Fig. 4).

The average UC associated with each LRC is computed when 1 and 2 information nodes update. In the following equations, we denote the average UC of our proposed LRC and that of other LRC by \overline{u}_x^{LRC1} and \overline{u}_x^{LRC2} , respectively. For u_1 , we have

$$\begin{aligned} 4 &\leq u_1 \leq 5, \\ \overline{u}_1^{LRC1} &= \frac{9(d-1)}{9} = d - 1 = 4, \text{ and} \\ \overline{u}_1^{LRC2} &= \frac{(d-1) + 8d}{9} = d - 0.11 = 4.89. \end{aligned}$$

Similarly, for u_2 , we have

$$\begin{aligned} 5 &\leq u_2 \leq 6, \\ \overline{u}_2^{LRC1} &= (26 \times 5 + 10 \times 6) / \binom{9}{2} = 5.27, \text{ and} \\ \overline{u}_2^{LRC2} &= (20 \times 5 + 16 \times 6) / \binom{9}{2} = 5.44. \end{aligned}$$

Hence, our proposed optimal LRC suggests UC close to the lower bound obtained in Theorem 1. In this case, our proposed optimal LRC improves the UC by 18.2% and

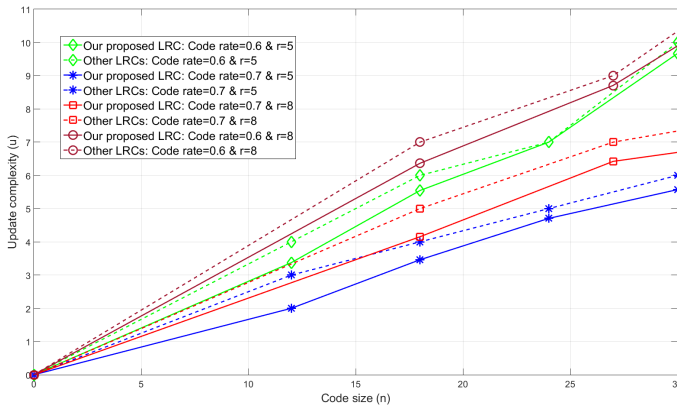


Figure 5: Comparison between the average UC of our proposed LRCs and other LRCs for two different code rate and code locality.

3.1% for \bar{u}_1 and \bar{u}_2 , respectively. Comparing with other approaches of designing optimal LRCs, this improvement is obtained without sacrificing other important properties of LRCs.

In Fig. 5, we compare the general approach of LRCs and our proposed optimal LRCs in terms of two different code rates and code localities.

V. CONCLUSION

The class of locally repairable codes (LRCs) is an important class of erasure codes to store data efficiently in distributed storage systems. In this paper, we established bounds on the update complexity (UC) of an important class of LRCs. Furthermore, we proposed a class of LRCs with small UC. Considering the recent usage of LRCs in practice, e.g. in Facebook HDFS-RAID and Windows Azure Storage, our proposed LRCs could also be of interest from a practical point of view.

REFERENCES

- [1] M. Sathiamoorthy, M. Asteris, D. Papailiopoulos, A. G. Dimakis, R. Vadali, S. Chen, and D. Borthakur, “Xoring elephants: Novel erasure codes for big data,” *Proc. VLDB*, vol. 6, no. 5, pp. 325–336, 2013.
- [2] C. Huang, H. Simitci, Y. Xu, A. Ogus, B. Calder, P. Gopalan, J. Li, S. Yekhanin et al., “Erasure coding in Windows Azure storage,” *Proc. USENIX Annual Technical Conference*, pp. 15–26, 2012.
- [3] D. Ford, F. Labelle, F. Popovici, M. Stokely, V.-A. Truong, L. Barroso, C. Grimes, and S. Quinlan, “Availability in globally distributed storage systems,” *Proc. USENIX Symposium on Operating Systems Design and Implementation*, pp. 61–74, 2010.
- [4] P. Gopalan, C. Huang, H. Simitci, and S. Yekhanin, “On the locality of codeword symbols,” *Information Theory, IEEE Transactions on*, vol. 58, no. 11, pp. 6925–6934, Nov 2012.
- [5] D. Papailiopoulos and A. Dimakis, “Locally repairable codes,” *Information Theory, IEEE Transactions on*, vol. 60, no. 10, pp. 5843–5855, Oct 2014.
- [6] A. Wang and Z. Zhang, “An integer programming-based bound for locally repairable codes,” *Information Theory, IEEE Transactions on*, vol. 61, no. 10, pp. 5280–5294, Oct 2015.
- [7] G. M. Kamath, N. Prakash, V. Lalitha, and P. V. Kumar, “Codes with local regeneration and erasure correction,” *IEEE Transactions on Information Theory*, vol. 60, no. 8, pp. 4637–4660, Aug 2014.
- [8] S. Goparaju and R. Calderbank, “Binary cyclic codes that are locally repairable,” in *Information Theory (ISIT), 2014 IEEE International Symposium on*, June 2014, pp. 676–680.

- [9] M. Shahabinejad, M. Khabbazi, and M. Ardakani, “An efficient binary locally repairable code for hadoop distributed file system,” *Communications Letters, IEEE*, vol. 18, no. 8, pp. 1287–1290, Aug 2014.
- [10] T. Ernvall, T. Westerback, R. Freij-Hollanti, and C. Hollanti, “Constructions and properties of linear locally repairable codes,” *IEEE Transactions on Information Theory*, vol. 62, no. 3, pp. 1129–1143, March 2016.
- [11] N. Silberstein, A. Rawat, O. Koyluoglu, and S. Vishwanath, “Optimal locally repairable codes via rank-metric codes,” pp. 1819–1823, July 2013.
- [12] I. Tamo and A. Barg, “A family of optimal locally recoverable codes,” *Information Theory, IEEE Transactions on*, vol. 60, no. 8, pp. 4661–4676, Aug 2014.
- [13] I. Tamo, D. Papailiopoulos, and A. Dimakis, “Optimal locally repairable codes and connections to matroid theory,” in *Information Theory Proceedings (ISIT), 2013 IEEE International Symposium on*, July 2013, pp. 1814–1818.
- [14] R. Tanner, “A recursive approach to low complexity codes,” *IEEE Transactions on Information Theory*, vol. 27, no. 5, pp. 533–547, Sep 1981.
- [15] F. R. Kschischang, B. J. Frey, and H. A. Loeliger, “Factor graphs and the sum-product algorithm,” *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 498–519, Feb 2001.
- [16] M. Shahabinejad, M. Khabbazi, and M. Ardakani, “A class of binary locally repairable codes,” *Communications, IEEE Transactions on*, accepted for publication, 2016.
- [17] N. P. Anthapadmanabhan, E. Soljanin, and S. Vishwanath, “Update-efficient codes for erasure correction,” in *Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on*, Sept 2010, pp. 376–382.
- [18] M. Mehrabi, M. Shahabinejad, M. Ardakani, and M. Khabbazi, “Optimal locally repairable codes with improved update complexity,” vol. abs/1606.09463, 2016. [Online]. Available: <http://arxiv.org/abs/1606.09463>
- [19] M. Shahabinejad, M. Ardakani, and M. Khabbazi, “An erasure code with reduced average locality for cloud storage systems,” *globeecom 2016, Under Review*, 2016.

APPENDIX A PROOF OF THEOREM 1

Here, we prove Theorem 1. First, we state a lemma to find the minimum parity nodes required to be connected to each information node.

Lemma 5. In the Tanner graph associated with an (n, k, d, r) ν -LRC shown in Fig. 2, any information node of non-infomixed and infomixed groups is linked to at least $d - 1$ and exactly $d - 1$ parity nodes, respectively.

Proof: The generator matrix associated with the ν -LRC related to Fig. 2 is

$$\mathbf{G} = [\mathbf{I}_k, \mathbf{P}] \in \mathbb{F}_q^{k \times n},$$

where $\mathbf{P} = [\mathbf{P}_1, \mathbf{P}_2] \in \mathbb{F}_q^{k \times (n-k)}$ is the parity matrix generator. Matrix $\mathbf{P}_1 \in \mathbb{F}_q^{k \times (\frac{n}{r+1} - \lambda)}$, which generates the local parity nodes of the first $\frac{n}{r+1} - \lambda$ local groups, can be presented as

$$\mathbf{P}_1 = \begin{bmatrix} \mathbf{I}_{\frac{n}{r+1} - \lambda} \otimes \mathbf{1}_r \\ \mathbf{0}_{(r-d+2 + \lfloor \frac{d-2}{r+1} \rfloor + (\lambda-1)r) \times (\frac{n}{r+1} - \lambda)} \end{bmatrix} \in \mathbb{F}_q^{k \times (\frac{n}{r+1} - \lambda)}.$$

As well, matrix $\mathbf{P}_2 \in \mathbb{F}_q^{k \times (d-2 - \lfloor \frac{d-2}{r+1} \rfloor + \lambda)}$ generates the local parity nodes of the λ mixed groups, and the global parity nodes. Observe that the i -th row of \mathbf{P}_1 has one non-zero element for $i \in [k - (r\lambda - d + 2 + \lfloor \frac{d-2}{r+1} \rfloor)]$ and the last

$r\lambda - d + 2 + \lfloor \frac{d-2}{r+1} \rfloor$ rows of \mathbf{P}_1 are all zero. In order to satisfy the minimum distance constraint, each row of \mathbf{G} must have at least d non-zero elements. Considering the first $k - (r\lambda - d + 2 + \lfloor \frac{d-2}{r+1} \rfloor)$ rows of \mathbf{G} , each row has two nonzero elements from \mathbf{I}_k and \mathbf{P}_1 . Thus, each row of \mathbf{P}_2 must have at least $d-2$ non-zero elements. In other words, each row of \mathbf{P}_2 has at most one zero element. Note that the last $r\lambda - d + 2 + \lfloor \frac{d-2}{r+1} \rfloor$ rows of \mathbf{P}_1 have no non-zero elements. Hence, at least $d-1$ elements of the last $r\lambda - d + 2 + \lfloor \frac{d-2}{r+1} \rfloor$ rows of \mathbf{P}_2 are non-zero. Note that if $r\lambda - d + 2 + \lfloor \frac{d-2}{r+1} \rfloor > 0$, i.e. there exists an infomixed group, we have $d-2 - \lfloor \frac{d-2}{r+1} \rfloor + \lambda = d-1$. Thus, \mathbf{P}_2 has $d-1$ columns; and all $d-1$ elements of the last $r\lambda - d + 2 + \lfloor \frac{d-2}{r+1} \rfloor$ rows of \mathbf{P}_2 are non-zero. ■

Proof of Lower bound:

According to Lemma 5, each row of \mathbf{P} has at least $d-1$ non-zero elements. Thus, any information node update leads to at least $d-1$ parity node updates. Regarding Fig. 2, the last $r\lambda - d + 2 + \lfloor \frac{d-2}{r+1} \rfloor$ information nodes are involved in the λ local parity nodes and $d-2 - \lfloor \frac{d-2}{r+1} \rfloor$ global parity nodes. Hence, information nodes associated with the infomixed group require exactly $d-1$ parity node updates which is the minimum required updates. Now, assume that the number of information nodes to be updated is less than or equal to the number of information nodes associated with the infomixed group, i.e. $x \leq r\lambda - d + 2 + \lfloor \frac{d-2}{r+1} \rfloor$. Then, assuming all the x information nodes are in the infomixed group, the lower bound on UC associated with x information node updates is $(d-1)$. On the other hand, if the number of information nodes to be updated exceeds the number of information nodes associated with the infomixed group, i.e. if $x > r\lambda - d + 2 + \lfloor \frac{d-2}{r+1} \rfloor$, we assume that $r\lambda - d + 2 + \lfloor \frac{d-2}{r+1} \rfloor$ information nodes are in the infomixed group and the remaining $x - r\lambda - d + 2 + \lfloor \frac{d-2}{r+1} \rfloor$ information nodes are in other local groups. Note that if an information node associated with the first $\frac{n}{r+1} - \lambda$ local groups is updated, then its local parity node has to be updated too. Thus, in this case, the total number of local parity nodes which have to be updated is $\lceil \frac{x - (r\lambda - d + 2 + \lfloor \frac{d-2}{r+1} \rfloor)}{r} \rceil$ and the lower bound on u_x is obtained as $u_x \geq d-1 + \lceil \frac{x - (r\lambda - d + 2 + \lfloor \frac{d-2}{r+1} \rfloor)}{r} \rceil$.

Proof of Upper bound:

In order to obtain the upper bound on u_x , we assume that x information nodes are located in distinct x local groups. From Lemma 5, any information node update requires at least $d-1$ parity node updates. Assuming that $x \leq \frac{n}{r+1} - \lambda$, among all x information node updates, there exists only one information node update in each local group. Thus, in this case, the number of local parity node updates is exactly x . On the other hand, if $x > \frac{n}{r+1} - \lambda$, each of the first $\frac{n}{r+1} - \lambda$ local groups has exactly one local parity node update. Hence, the upper bound on u_x is obtained as $u_x \leq d-1 + \theta$, where $\theta = x$ if $x \leq \frac{n}{r+1} - \lambda$ and $\theta = \frac{n}{r+1} - \lambda$ otherwise. Observe that for $x > \frac{n}{r+1} - \lambda$,

$d-1 + \theta = n - k$ which is the amount of all parity nodes.

APPENDIX B PROOF OF THEOREM 2

Here, we prove Theorem 2. By Theorem 1, we know that updating an information node requires updating either $d-1$ or d parity nodes. In the following, first, we state a lemma to find the minimum number of information nodes connected to d parity nodes in any set of $\lfloor \frac{d-2}{r+1} \rfloor + 1$ non-mixed groups. Then, by using this lemma, we prove the established lower and upper bounds.

Lemma 6. Within any collection of $\lfloor \frac{d-2}{r+1} \rfloor + 1$ local groups in an (n, k, d, r) ν -LRC shown in Fig. 2, the minimum number of information nodes except the information nodes of the infomixed group, denoted η , with UC equal to exactly d is

$$\eta = \begin{cases} 0 & \text{if } \frac{\alpha}{(d-2 - \lfloor \frac{d-2}{r+1} \rfloor)} \mid \beta \text{ or } \lfloor \frac{d-2}{r+1} \rfloor = \lceil \frac{d-2}{r+1} \rceil, \\ \alpha - \beta \lfloor \frac{\alpha}{\beta} \rfloor & \text{otherwise} \end{cases},$$

where $\alpha = ((r+1)(\lfloor \frac{d-2}{r+1} \rfloor + 1) - (d-2))(d-2 - \lfloor \frac{d-2}{r+1} \rfloor)$ and $\beta = r(\lfloor \frac{d-2}{r+1} \rfloor + 1)$.

Proof: In an (n, k, d, r) ν -LRC, there are $d-2 - \lfloor \frac{d-2}{r+1} \rfloor$ global check nodes. By Theorem 3, a necessary condition to guarantee a minimum distance d for our proposed codes is that any collection of check nodes consisting of $n - k - d + 1$ local check nodes and a single global check node must cover at least $n - (d-2)$ variable nodes. Hence, each of the global check nodes have to be connected to at least $(r+1)(\lfloor \frac{d-2}{r+1} \rfloor + 1) - (d-2)$ information nodes in any set of $\lfloor \frac{d-2}{r+1} \rfloor + 1$ local groups. Consequently, for every set of $\lfloor \frac{d-2}{r+1} \rfloor + 1$ local groups, there exist at least $\alpha = ((r+1)(\lfloor \frac{d-2}{r+1} \rfloor + 1) - (d-2))(d-2 - \lfloor \frac{d-2}{r+1} \rfloor)$ connections between the information nodes and the global check nodes. Also, there exist r information nodes in each local group but the mixed groups. Thus, in a set of $\lfloor \frac{d-2}{r+1} \rfloor + 1$ non-mixed groups, we have $\beta = r(\lfloor \frac{d-2}{r+1} \rfloor + 1)$ information nodes. Consequently, in a set of $\lfloor \frac{d-2}{r+1} \rfloor + 1$ non-mixed groups, each information node has to be connected to at least $\lfloor \frac{\alpha}{\beta} \rfloor$ global check nodes; hence, for all the information nodes in this set, we need at least $\beta \lfloor \frac{\alpha}{\beta} \rfloor$ connections to global check nodes. Therefore, by subtracting $\beta \lfloor \frac{\alpha}{\beta} \rfloor$ from all the α connections, we obtain the minimum number of information nodes connected to d parity nodes (η) in a set of $\lfloor \frac{d-2}{r+1} \rfloor + 1$ non-mixed groups as

$$\eta = \alpha - \beta \lfloor \frac{\alpha}{\beta} \rfloor.$$

Also, if $\lfloor \frac{d-2}{r+1} \rfloor = \lceil \frac{d-2}{r+1} \rceil$, there is no infomixed group; and all the information nodes are connected to exactly $d-1$ parity nodes, and consequently, $\eta = 0$. Furthermore, by Remark 5, if $\frac{\alpha}{(d-2 - \lfloor \frac{d-2}{r+1} \rfloor)} = ((r+1)(\lfloor \frac{d-2}{r+1} \rfloor + 1) - (d-2)) \mid \beta$, the same as the previous case, all the information nodes are connected to exactly $d-1$ parity nodes, and consequently, $\eta = 0$. ■

By Lemma 6, the minimum total number of information nodes connected to exactly d parity nodes is equal to $\eta \left\lfloor \frac{\frac{n}{r+1} - \lambda}{\lfloor \frac{d-2}{r+1} \rfloor + 1} \right\rfloor$. Hence, we have

$$\frac{\eta \left\lfloor \frac{\frac{n}{r+1} - \lambda}{\lfloor \frac{d-2}{r+1} \rfloor + 1} \right\rfloor d + (k - \eta \left\lfloor \frac{\frac{n}{r+1} - \lambda}{\lfloor \frac{d-2}{r+1} \rfloor + 1} \right\rfloor)(d-1)}{k} \leq \overline{u_1} \leq \frac{kd}{k}.$$

Therefore,

$$(d-1) + \frac{\eta \left\lfloor \frac{\frac{n}{r+1} - \lambda}{\lfloor \frac{d-2}{r+1} \rfloor + 1} \right\rfloor}{k} \leq \overline{u_1} \leq d.$$

APPENDIX C

PROOF OF PROPOSITION 4

Here, we prove Proposition 4. For an (n, k) erasure code with Tanner graph \mathcal{T} , if any φ variable nodes are

connected to φ distinct check nodes, where $\varphi \in [d-1]$, then any φ variable nodes can be recovered using equations associated with the distinct check nodes. This implies that the code can recover up to any $d-1$ failures and therefore, it has minimum distance d .

In our proposed LRCs, every information nodes is connected to at least $(d-2 - \lfloor \frac{d-2}{r+1} \rfloor - (\lambda - \lfloor \frac{d-2}{r+1} \rfloor) + \lambda) + 1 = d-1$ distinct check nodes, where $\lambda = \left\lceil \frac{d-2 - \lfloor \frac{d-2}{r+1} \rfloor}{r} \right\rceil$. Also, each of the $n-k$ parity nodes is connected to exactly one distinct check node. Hence, any $d-1$ variable nodes are connected to at least $d-1$ distinct check nodes and our proposed code has minimum distance d .